Al-Farabi Kazakh National University (KazNU)

Faculty of Biology and Biotechnology

Department of Molecular Biology and Genetics

Discipline: «Modern Problems of Plant Genetics»

Lecture 1

Modern Problems of Plant Genetics

From classical crop genetics and the Green Revolution to T2T pangenomics, multi-omics integration, and genomic selection

Author: Amirova Aigul Kuzembayevna, Candidate of Biological Sciences, Associate Professor
E-mail: aigul_amir@mail.ru

Aim and Structure of the Lecture

Aim: introduce contemporary challenges in biology and plant genetics, high-throughput sequencing, and molecular breeding paradigms

01 Grand Challenges of Biology

Climate change, food security, biodiversity conservation, and decoding genome-to-phenome rules

02 Evolution of Plant Breeding

Transition from phenotypic mass selection and the Green Revolution to precision breeding (Breeding 4.0)

03 Sequencing Revolution (NGS)

From Sanger sequencing to Illumina short reads, PacBio HiFi, and Oxford Nanopore ultralong reads

04 Plant Genome Milestones

From Arabidopsis thaliana (2000) to telomere-to-telomere (T2T) crop assemblies and graph pangenomes

05 Multi-Omics & Functional Genetics

Integrating genomics, transcriptomics, proteomics, metabolomics, and high-throughput phenomics

06 MAS & Genomic Selection

Bridging genotype and complex agronomic traits using GWAS, QTL mapping, and GEBV prediction

Core Frontiers of Modern Biology in the 21st Century

Systemic questions uniting molecular genetics, biotechnology, ecology, and agriculture

01 Global Climate & Abiotic Stress

- Rising temperatures, erratic precipitation, soil salinization, and drought threaten major cereal yields
- Need to dissect polygenic stress-tolerance networks and root system architecture

02 Global Food & Nutrition Security

- Projected global population ~9.7 billion by 2050 requires +50–60% food production on finite arable land
- Simultaneous demand to eliminate micronutrient deficiencies («hidden hunger»)

03 Genotype-to-Phenotype (G2P) Gap

- While DNA sequencing has become fast and inexpensive, predicting complex field phenotypes remains challenging
- Epistasis, gene-by-environment interactions (G×E), and epigenetic regulation drive phenotypic plasticity

04 Genetic Erosion & Domestication Bottlenecks

- Modern elite cultivars retain only a fraction of the allelic diversity present in landraces and wild relatives
- Urgent need to characterize and mobilize germplasm collections via pre-breeding

Four Eras of Plant Genetics and Crop Improvement

Paradigm shifts from domestication to genomics-assisted design (Wallace et al.)

Breeding 1.0

Domestication (~10,000 BP)

- Unconscious mass selection of non-shattering spikes and larger seeds
- Founder bottlenecks in wheat, rice, maize

Breeding 2.0

Mendelian & Hybrid Era (1900–1960s)

- Rediscovery of Mendel's laws & heterosis
- Green Revolution: semi-dwarf wheat (Rht) & rice (sd1)

Breeding 3.0

Biotech & Molecular Markers (1980–2010)

- Transgenic GM crops & tissue culture
- QTL mapping and Marker-Assisted Selection (MAS)

Breeding 4.0

Precision Design (2015–2026)

- CRISPR/Cas genome editing & pangenomics
- Genomic Selection (GS), speed breeding & AI models

Two Complementary Strategies for Functional Plant Gene Discovery

Bridging observable agronomic variation and underlying nucleotide sequence

Forward Genetics («Phenotype → Genotype»)

- Starts with natural allelic variation or induced mutant populations (EMS, fast neutrons, T-DNA insertion)
- Screens thousands of lines for a visible phenotype (disease resistance, flowering time, dwarfism)
- Identifies the causal locus via biparental QTL linkage mapping or Genome-Wide Association Studies (GWAS)
- Gold standard for uncovering completely uncharacterized pathways and novel agronomic genes

↔

Reverse Genetics («Genotype → Phenotype»)

- Starts with a known candidate DNA sequence identified by homology or RNA-seq expression profiles
- Perturbs gene function directly using CRISPR/Cas9 knockout, TILLING, RNA interference (RNAi), or overexpression
- Quantifies resulting physiological, metabolic, or developmental changes under controlled growth conditions
- Rapidly validates candidate genes mined from QTL intervals and pangenome datasets

Evolution of Sequencing Technologies Driving Plant Genomics

Why third-generation long reads solved the repetitive-DNA assembly bottleneck in crops

Generation / Platform	Read Length	Accuracy & Throughput	Key Application in Plant Genetics
1st Gen: Sanger (Chain Termination)	600–1,000 bp	99.99%; very low throughput	Targeted PCR amplicon validation and clone-by-clone BAC sequencing (Arabidopsis 2000)
2nd Gen: Illumina SBS (Short Reads)	150–300 bp (paired)	>99.9%; ultra-high throughput, lowest $/Gb	Whole-genome resequencing, GBS, GWAS panels, RNA-seq, epigenomics
3rd Gen: PacBio HiFi (CCS)	15–25 kb	>99.9% per read (Q30+); high contiguity	De novo telomere-to-telomere (T2T) assembly of polyploid wheat, potato, and pangenomes
3rd Gen: Oxford Nanopore (ONT)	20 kb to >2 Mb	98–99% (Q20+); native 5mC methylation	Spanning megabase centromeric arrays, direct RNA sequencing, rapid in-field pathogen ID

Growth of Sequenced Plant Genomes and Polyploid Milestones (2000–2026)

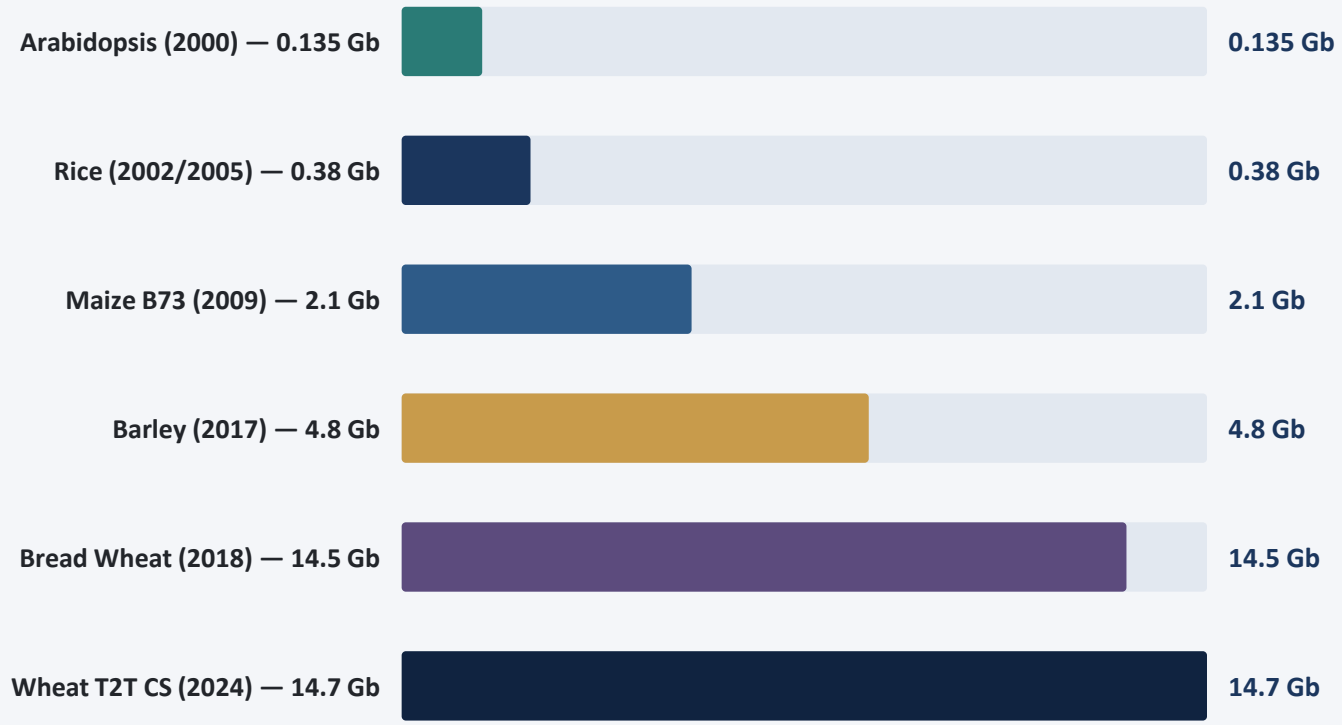
Increasing plant genome size (Mbp) tackled from early model species to gapless hexaploid wheat and oat assemblies

2000
Arabidopsis thaliana (1st plant)

>1,500
Plant genomes assembled

14.5 Gb
Hexaploid wheat (AABBDD)

Reference Assembly Sizes Across Landmark Plant Species (Gbp scale, log-representative)



From Single References to Pangenomes

- Early short-read assemblies fragmented at LTR retrotransposons covering 70–85% of cereal genomes
- In 2024–2025, gapless T2T assemblies of wheat, rice, maize, and soybean resolved centromeres and rDNA repeats
- Graph pangenomes capture structural variants (SVs) missed by a single reference cultivar

Multi-Omics Architecture for Dissecting Complex Agronomic Traits

Tracing biological information flow from DNA sequence variations to field-level yield and stress resilience

1 1. Genomics & Epigenomics

SNPs, InDels, structural variants (SVs), DNA methylation (5mC), histone marks

2 2. Transcriptomics

Bulk, tissue-specific, and single-cell RNA-seq capturing spatio-temporal gene expression

3 3. Proteomics

LC-MS/MS quantification of enzymes, signaling complexes, and post-translational modifications

4 4. Metabolomics

Primary sugars/amino acids and specialized secondary metabolites (flavonoids, alkaloids)

5 5. High-Throughput Phenomics

UAV multispectral drones, LIDAR, thermal imaging, and automated root phenotyping

Machine learning and network modeling integrate multi-omics layers to overcome the G2P bottleneck in modern crop improvement.

Modern Breeding Tools: Marker-Assisted Selection vs Genomic Selection

Matching breeding strategy to the genetic architecture of simple (oligogenic) vs complex (polygenic) traits

Marker-Assisted Selection (MAS)

- Best suited for major-effect Mendelian loci or large-effect QTLs (e.g., rust resistance genes Yr/Lr/Sr, Sub1A flood tolerance)
- Tracks 1 to 10 diagnostic DNA markers (KASP, SSR, CAPS) co-segregating with the target allele
- Ideal for rapid marker-assisted backcrossing (MABC) to introgress resistance from landraces without linkage drag
- Limitation: misses hundreds of small-effect polygenes controlling grain yield and drought resilience

↔

Genomic Selection (GS — Meuwissen et al.)

- Uses thousands of genome-wide SNPs simultaneously across a phenotyped Training Population to calibrate a statistical/AI model
- Computes Genomic Estimated Breeding Values (GEBV) for unphenotyped Selection Candidates at the seedling stage
- Captures both major and minor polygenic effects simultaneously without prior QTL significance thresholding
- Combined with Speed Breeding (4–6 generations/year in wheat/barley), doubles genetic gain per unit time (ΔG)

Priority Research Pillars in Contemporary Plant Genetics

Connecting fundamental chromosomal and molecular discoveries to agricultural resilience in Kazakhstan and globally

Pillar 1 Climate-Resilient Ideotypes

- Dissecting heat, drought, frost, and salinity tolerance networks in cereals and legumes
- Optimizing flowering time genes (Vrn, Ppd) and root water-uptake efficiency

Pillar 2 Durable Biotic Resistance

- Cloning NLR immune receptors and susceptibility (S) genes against rusts, Fusarium, and viruses
- Gene pyramiding and CRISPR knockout of S-genes (e.g., MLO, SWEET) for broad-spectrum durability

Pillar 3 Photosynthetic & Nitrogen Efficiency

- Enhancing Rubisco kinetics, photorespiratory bypasses, and C4-like leaf anatomy in C3 crops
- Reducing synthetic fertilizer dependence via microbiome symbiosis and NUE alleles

Pillar 4 Central Asian Agrobiodiversity

- Genomic characterization of Kazakhstani wheat cultivars, Malus sieversii, and wild Hordeum/Aegilops
- Utilizing doubled haploids, somatic hybrids, and alien chromosome translocations

Accelerating Genetic Gain: Integrating Speed Breeding and Genomic Selection

Reducing wheat and barley cultivar development cycles from 10–12 years down to 3–4 years

1 1. Strategic Crossing

Select complementary parents using pangenome haplotype diversity

2 2. Rapid Generations (F1–F4)

22-hour LED photoperiod accelerates wheat/barley up to 6 generations per year

3 3. Seedling Genotyping

High-density SNP array or low-pass sequencing predicts GEBV before flowering

4 4. Doubled Haploids / Editing

Immediate fixation of homozygosity or targeted multiplex CRISPR modification

5 5. Multi-Env Field Trials

Verify yield stability and grain quality across Kazakhstan agro-ecological zones

Course Architecture: Key Themes Covered in Lectures 1–10

How classical chromosome biology, cell culture biotechnology, microbiome ecology, and genome editing integrate

Module	Lectures	Core Genetic & Biotechnological Focus
Genomic Foundations	Lectures 1–2	Modern problems, NGS/multi-omics, C-value paradox, TEs, polyploidy, and pangenomics
Cellular Biotechnology	Lectures 3–4	In vitro culture, anther/microspore doubled haploids, protoplast fusion & somatic hybrids
Germplasm & Chromosomes	Lecture 5	Plant Genetic Resources (Vavilov, Kazakhstan wild apple), aneuploids & alien translocations
Biosafety, Ecology & Law	Lectures 6–8	ERA of GM/edited crops, plant microbiome & PGPB, global agritech market & regulations
Impact & Precision Editing	Lectures 9–10	UN SDG 2 Zero Hunger biofortification & CRISPR/Cas9, Base/Prime Editing in agriculture

Translating Plant Genetics from Laboratory to Field

Scientific breakthroughs must align with ecological safety, international trade treaties, and public trust

Science-Based Risk Assessment

- Evaluating gene flow to wild relatives, non-target organism effects, and food allergenicity
- Distinguishing SDN-1 transgene-free gene edits from classical transgenic GMOs

Sustainable Agroecosystems

- Harnessing plant-growth-promoting bacteria (PGPB) and mycorrhizae instead of excessive agrochemicals
- Protecting beneficial pollinators and rhizosphere biodiversity

PGRFA & Access and Benefit Sharing

- Balancing breeders' intellectual property (UPOV) with equitable germplasm exchange (ITPGRFA)
- Empowering smallholder farmers with resilient, non-proprietary seed options

Evidence-Based Public Policy

- Transparent labeling, traceability, and clear dialogue addressing public misconceptions
- Harmonizing national biosafety laws in Kazakhstan with international Codex and FAO standards

The Breeder's Equation: Quantifying Progress in Modern Plant Genetics

How genomic tools directly optimize every term of genetic gain per year: $\Delta G = (i \times r \times \sigma_A) / L$

Components of the Breeder's Equation (ΔG)

- Selection intensity (i): massively expanded by automated seedling genotyping and drone phenotyping
- Selection accuracy (r): improved for low-heritability traits (drought yield) using multi-trait genomic prediction
- Additive genetic variance (σ_A): replenished via wild relative introgressions and de novo domestication
- Breeding cycle length (L): cut from 8–10 years to 1–2 years through speed breeding and doubled haploids

↔

Urgent Targets for 2030–2050 Crops

- Maintain or increase calorie and protein output under +1.5 to +2.0 °C warming scenarios
- Achieve 20–30% lower synthetic nitrogen and water footprint per ton of harvested grain
- Combat emerging transboundary crop pathogens (wheat stem rust Ug99, Fusarium, bacterial blights)
- Deploy nutritional biofortification (Fe, Zn, provitamin A, folate) directly in staple crop varieties

Control Questions and Key Takeaways

Consolidating Lecture 1 concepts

? Control & Review Questions

- 1. What are the four major challenges facing modern biology and plant genetics in the 21st century?
- 2. Contrast the eras of Breeding 1.0 through Breeding 4.0 and define the Genotype-to-Phenotype (G2P) gap.
- 3. Compare Forward Genetics and Reverse Genetics approaches for uncovering plant gene function.
- 4. Why did third-generation long-read sequencing (PacBio HiFi, ONT) transform plant genomics compared to short reads?
- 5. Distinguish Marker-Assisted Selection (MAS) from Genomic Selection (GS) and explain how Speed Breeding accelerates ΔG.

✓ Key Takeaways

- Modern plant genetics integrates genomics, multi-omics, phenomics, and AI to breed climate-resilient, nutrient-dense crops.
- Long-read HiFi/ONT technologies unlocked gapless T2T genomes and graph pangenomes even in giant polyploid cereals.
- MAS targets major oligogenic loci, whereas Genomic Selection (GEBV) models polygenic traits across genome-wide SNPs.
- Accelerating genetic gain (ΔG) requires combining germplasm diversity with doubled haploids, speed breeding, and gene editing.

Recommended Literature and Online Databases

Core textbooks, reviews, and genomic portals for further study

1 Main Literature

- 1. Wallace J.G., Rodgers-Melnick E., Buckler E.S. On the Road to Breeding 4.0: Unraveling the Good, the Bad, and the Boring of Crop Quantitative Genomics // Annu. Rev. Genet. — 2018. — Vol. 52.
- 2. Michael T.P., VanBuren R. Building near-complete plant genomes // Current Opinion in Plant Biology. — 2020. — Vol. 54. — P. 26–33.
- 3. Hickey L.T. et al. Breeding crops to feed 10 billion // Nature Biotechnology. — 2019. — Vol. 37. — P. 744–754.
- 4. Brown T.A. Genomes 5. — Garland Science / CRC Press, 2023.

2 Additional Literature & Databases

- 1. Ensembl Plants (comparative plant genomics portal): https://plants.ensembl.org
- 2. Phytozome (JGI Comparative Plant Genomics): https://phytozome-next.jgi.doe.gov
- 3. Gramene (cereal functional genomics & pathways): https://www.gramene.org
- 4. FAO State of Food and Agriculture Reports: https://www.fao.org/publications

Thank you for your attention! • Department of Molecular Biology and Genetics, KazNU